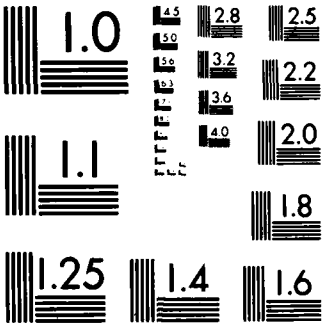
AD-A131 617 DETERMINING ATTITUDE OF OBJECT FROM NEEDLE MAP USING EXTENDED GAUSSIAN IMAGE(U) MASSACHUSETTS INST OF TECH CAMBRIDGE ARTIFICIAL INTELLIGENCE L.. K IKEUCHI APR 83 1/1

UNCLASSIFIED AI-M-714 N00014-77-C-0389 F/G 9/2 NL



UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AI-Memo-714	2. GOVT ACCESSION NO. AD-A131617	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Determining Attitude of Object From Needle Map Using Extended Gaussian Image		5. TYPE OF REPORT & PERIOD COVERED 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Katsushi Ikeuchi		8. CONTRACT OR GRANT NUMBER(s) N00014-77-C-0389 N00014-80-C-0505
9. PERFORMING ORGANIZATION NAME AND ADDRESS Artificial Intelligence Laboratory 545 Technology Square Cambridge, Massachusetts 02139		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Advanced Research Projects Agency 1400 Wilson Blvd Arlington, Virginia 22209		12. REPORT DATE April 1983 13. NUMBER OF PAGES 26
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Office of Naval Research Information Systems Arlington, Virginia 22217		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this document is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Computer Vision
Object Attitude
Bin Picking
Lookup table
Gauss Map
least Inertia Axis
Tessellated Dome

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An extended Gaussian image (EGI) is constructed by mapping the surface normals of an object onto the Gaussian sphere. The attitude of an object is greatly constrained by the global distribution of EGI mass over the visible Gaussian hemisphere. Constraints on the viewer direction are derived from the position of the EGI mass center, and from the diection of the EGI inertia axis. The algorithm embodying these constraints and the EGI mass distribution are implemented using a lookup table. A function for matching an observed EGI with the prototypical EGIs is also proposed, The algorithm determines the attitude of an object over

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

successfully both from a synthesized needle map and a real needle map.

Massachusetts Institute of Technology

Artificial Intelligence Laboratory

A.I. Memo No. 714 April, 1983

Determining Attitude of Object From Needle Map Using Extended Gaussian Image

Katsushi Ikeuchi

Abstract

An extended Gaussian image (EGI) is constructed by mapping the surface normals of an object onto the Gaussian sphere. The attitude of an object is greatly constrained by the global distribution of EGI mass over the visible Gaussian hemisphere. Constraints on the viewer direction are derived from the position of the EGI mass center, and from the direction of the EGI inertia axis. The algorithm embodying these constraints and the EGI mass distribution are implemented using a lookup table. A function for matching an observed EGI with the prototypical EGIs is also proposed. The algorithm determines the attitude of an object successfully both from a synthesized needle map and a real needle map .

Key Words: Computer Vision, Hand-Eye System, Object Attitude, Bin Picking, Gauss Map, Least Inertia Axis, Tessellated Dome, Lookup table

Acknowledgment: This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Office of Naval Research under Office of Naval Research contract N00014-77-C-0389 and in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-80-C-0505.

1. Introduction

One of the most important tasks for the visual guidance of manipulators is the determination of the attitude of an object relative to a camera. This attitude information determines the direction from which an end effector should approach and grasp the object. If the object lies on a horizontal table, the gravity direction may constrain this attitude [1]. In the general case, however, one has to determine the attitude by comparing information provided by a camera with some internal model of the object.

Historically, this attitude-determination problem has been solved by comparing an observed silhouette with some internal shape description [2-5]. These edge-based approaches work well in determining the attitude of an isolated object lying on a uniform background provided the object is able to rotate only in the plane of support. In other words, these algorithms work well on binary images. However, edge-based methods have difficulties in extracting the contour of a particular object from a set of many overlapping objects, which are typical in a bin-picking problem.

Recent work in Image Understanding [6-9] has let to techniques for computing local surface gradient by various means. Such methods include: shape from shading [10-13], photometric stereo [14-16], shape from contours, and shape from texture [17-28]. This local gradient representation is referred to as a needle map[22], or $2\frac{1}{2}D$ sketch of the scene [23-24]. Since this local information is obtained over a whole region within some boundaries, it is more robust than silhouette information which comes only from the boundaries. This paper focuses on the problem of determining the attitude of an object from this surface normal representation.

2. Needle map and Extended Gaussian image

A needle map is an assignment of surface normals to patches of an object corresponding to pixels in the image. A needle map is expressed in a viewer-centered coordinate system. Each surface patch is projected onto the image plane. Each surface normal is expressed with respect to the line of sight. Thus, a needle map describes surface orientation relative to the viewer. On the other hand, the usual

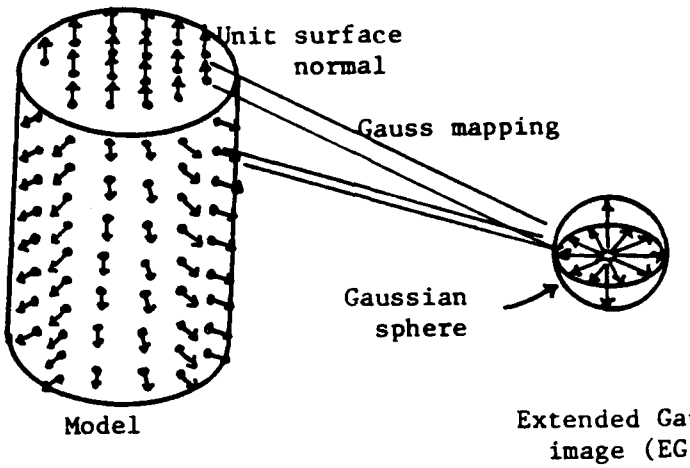


Figure 1. A spike model and EGI. The normals on a spike model can be moved to a common point of application. Then, the locus of their end points lies on the surface of a unit sphere. The distribution of these end points is called the extended Gaussian image.

---

internal model describes an object based on characteristics of the object such as its center of mass, and axis of least inertia. Such a method describes surface orientation relative to the object mass center and is called an object-centered coordinate system [22,24,25]. These two kinds of coordinate systems are independent of each other. Our purpose is to determine the transformation between them using an extended Gaussian image [22,26,28,29].

Roughly speaking, the extended Gaussian image of an object is a spatial histogram of its surface normals. Let us assume that there is a fixed number of surface patches per unit surface area and that a unit normal is erected on each patch. The collection of these normals, which are like porcupine's quills, is called a spike model of the object [22]. These normals can be moved to a common point of application so that the locus of their end points lies, then, on the surface of a unit sphere. This mapping is called the Gauss map; the unit sphere is called the Gaussian sphere [22] (See Fig 1). If we attach a unit mass to each end point, we will observe a distribution of mass over the Gaussian sphere. The resulting distribution of mass is called the extended Gaussian image (EGI) of the object [1,8].

Let us define a visible hemisphere. Commonly, one observes an object from one

direction. So we always obtain only one half of the EGI over a Gaussian hemisphere. This hemisphere will be referred to as the visible hemisphere. The pole of the visible hemisphere corresponds to the line of sight. Each point on the visible hemisphere corresponds to the surface orientation whose inner angle between the line of sight is no more than $\frac{\pi}{2}$. In the following discussion we will work with this EGI over the visible hemisphere. Also we will normalize the distribution of EGI mass to have a unit mass over the visible hemisphere.

Even though an EGI can be produced from a spike model based on the object-centered coordinate system, the EGI can be easily interpreted in the viewer-centered coordinate system. In this case, an apparent image of an object varies with the following factors: (a)translation of the object, (b)expansion of the object, (c)rotation of the object. The normalized EGI is independent of (a) and (b). The EGI rotates in the same way as (c) as will be shown next:

(a) Neither the surface normals nor the Gauss mapping depend on the position of the origin. Thus, the resulting EGI is not affected by translation of the object.

(b) If the object expands, the total mass over the Gaussian hemisphere increases. Yet, the EGI mass is normalized so as to have a unit mass over the hemisphere. Thus, the normalized EGI does not change upon object expansion. This characteristic is very convenient in object recognition. In general, the distance between the TV camera and the object changes in each situation. Thus, the apparent size of an object will also vary, although the normalized EGI derived from the image is independent of the apparent size.

(c) When an object rotates, its EGI also rotates. Fortunately, the EGI rotates in the same manner as the object. In other words, this rotation does not effect the relative EGI mass distribution over the sphere. This is analogous to the fact that the relative distribution of continents on the earth does not change as the earth rotates (See Fig. 2). Thus, if an observed EGI is identical to one part of the model's EGI, we can find which part of the object is observed at that time, thus we can find the object's relative attitude.

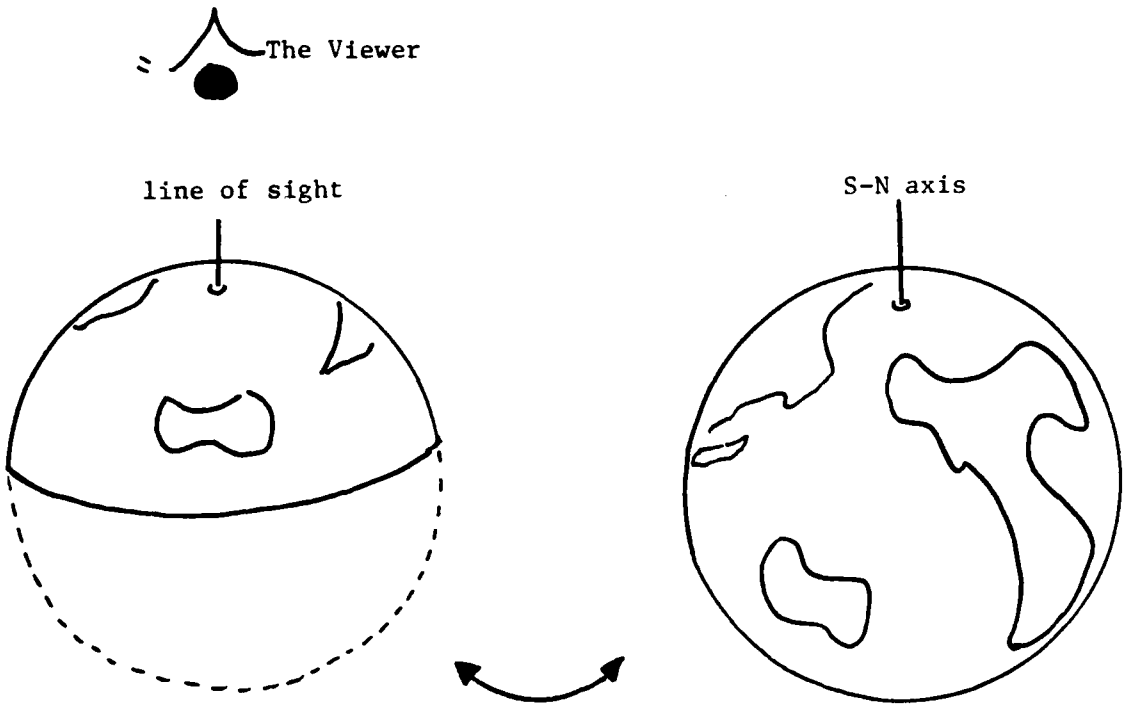


Figure 2. EGI rotation. When an object rotates, its EGI also rotates in the same manner as the object. If an observed EGI is identical to one part of the prototypical EGI, we can find which part of the object is observed.

3. Constraints from a global EGI distribution

Generally speaking, the apparent image depends on the transformation between the viewer-centered coordinate system and the object-centered coordinate system. This transformation has six degrees of freedom: three degrees of freedom in position and three degrees of freedom in attitude. From characteristics (a) and (b), an observed EGI does not depend on the position freedom. The EGI only depends on the attitude freedom. Thus, matching an observed EGI with a model EGI involves three degrees of freedom. There are two degrees of freedom corresponding to which point on the Gaussian sphere is perpendicular to the line of sight. The remaining degree of freedom comes from rotation about the line of sight.

We will use two constraints to reduce these degrees of freedom. Although a brute force technique, such as search through the space of possible attitudes, can be used [28,29], we will reduce this search space using constraints before EGI

comparison. The EGI mass center position constrains the line of sight. Furthermore, the EGI inertia direction constrains the rotation around the line of sight.

3.1. Coordinate Definition

The following discussion will refer three local coordinate systems: (1) a coordinate system attached to the prototype Gaussian sphere, (2) a coordinate system attached to the observed Gaussian hemisphere, and (3) a coordinate system on the observed image plane.

The Z axis of the prototype Gaussian sphere agrees with the S-N axis of the sphere, the X-Y plane corresponds to the cross section along the equator. These axes are denoted by X_p, Y_p, Z_p.

On the observed Gaussian hemisphere, the S-N axis corresponds to the line of sight direction. The X-Y plane is the base plane of the hemisphere. This coordinate system is denoted by X_o, Y_o, Z_o. The coordinate axes are rotated about the Z_p axis until X_o is perpendicular to the $Z_p - V$ plane, where V is the line of sight. This means that X_o, Y_o, Z_o is rotated around the Z_p axis by $(\phi + \frac{\pi}{2})$, where ϕ is the azimuth angle of the line of sight as shown in Fig. 3. Then, the coordinate system is rotated by θ around the X_o axis until the Z_o axis agrees with V, where θ is the zenith angle of V. Thus, the resulting coordinate transformation may be written as

$$X_o = -X_p \sin\phi + Y_p \cos\phi \tag{1.1}$$
$$Y_o = -X_p \cos\theta \cos\phi - Y_p \cos\theta \sin\phi + Z_p \sin\theta \tag{1.2}$$
$$Z_o = X_p \sin\theta \cos\phi + Y_p \sin\theta \sin\phi + Z_p \cos\theta. \tag{1.3}$$

Finally, the image plane is assumed to be tangent to the visible Gaussian hemisphere at the north pole. Its coordinates are denoted as X_i, Y_i. So,

$$X_i = X_o \tag{2.1}$$
$$Y_i = Y_o \tag{2.2}$$

3.2. EGI mass center

Elevation of the EGI mass center from the hemisphere base plane gives a constraint on the line of sight. Even though the EGI mass center over the

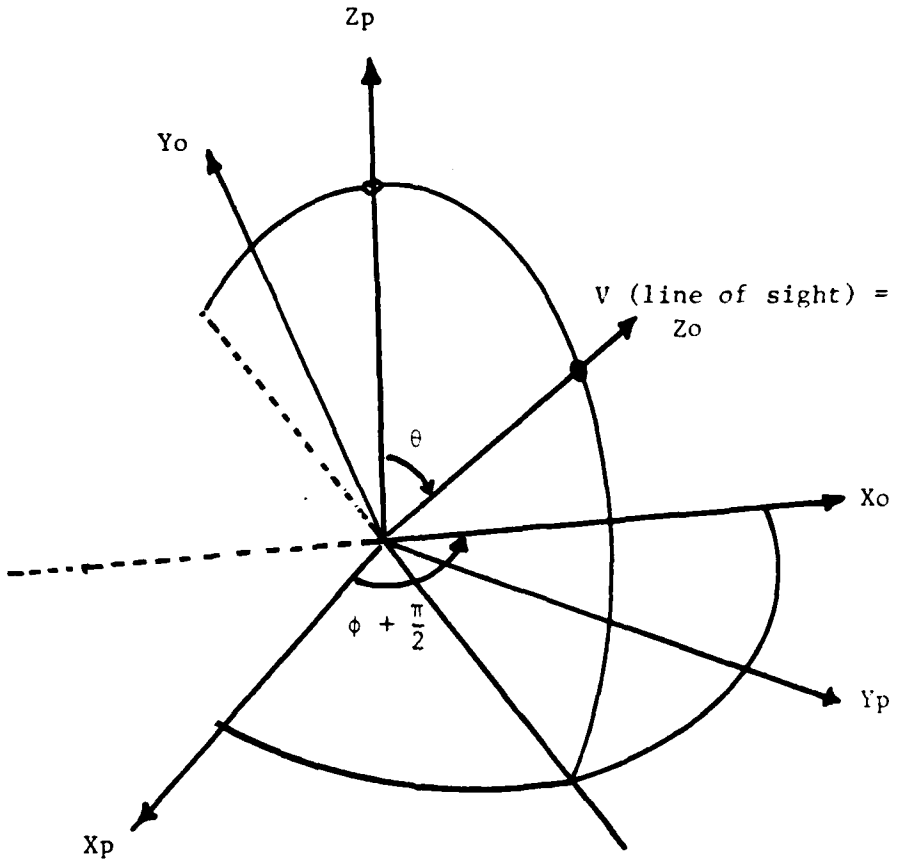


Figure 3. The relationship between two coordinate systems. X_o, Y_o, Z_o is rotated around the Z_p axis by $(\phi + \frac{\pi}{2})$, Then, the coordinate system is rotated by θ around the X_o axis until the Z_o axis agrees with the line of sight, V.

---

whole sphere is at the center of the sphere, the EGI distribution over a visible hemisphere always has some bias. Since this mass center is different for different visible hemispheres, correspondence of the EGI mass centers becomes a necessary condition for correspondence of the EGI distribution. Thus, comparing the observed EGI mass center with the prototype's reduces the freedom of the line of sight.

This elevation $A(v)$ at the direction, v is obtained as

$$A(v) = \frac{\int\int_{V.H.} Z_o(s,t) EGIM(s,t) \sqrt{EG - F^2} ds dt}{\int\int_{V.H.} EGIM(s,t) \sqrt{EG - F^2} ds dt}, \tag{3}$$

where (s,t) is a parameterization over the Gaussian hemisphere, $EGIM(s,t)$ is EGI mass there, and V.H. stands for a visible hemisphere of v. $(X_o(s,t), Y_o(s,t), Z_o(s,t))$

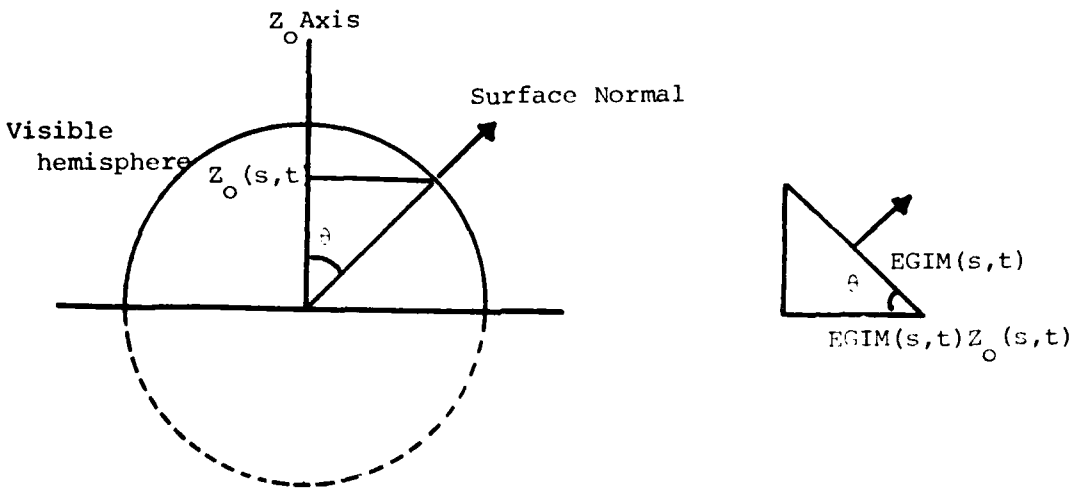


Figure 4. The visible hemisphere and the EGI mass. $Z_o(s,t)$ is equal to the cosine of the angle between the surface normal and the Z_o axis. Thus, $Z_o(s,t)EGIM(s,t)$ denotes the projected area.

---

denotes the coordinate value of the point (s,t) on the Gaussian hemisphere in the observed coordinate system X_o, Y_o, Z_o. E, F, G are the first fundamental form of the Gaussian hemisphere of the parametrization (s,t) [30, pg.268]. $\sqrt{EG - F^2}$ may be regarded as a jaccobian from (s,t) to the hemisphere surface. Note that the line of sight, v is equivalent to the Z_o axis from the definition.

We will call this elevation the projection area ratio, because this value equals the ratio of projected area to surface area. Since $EGIM(s,t)$ is equivalent to the surface area of patches whose normal is $(X_o(s,t), Y_o(s,t), Z_o(s,t))$, the denominator represents the total surface area. On the other hand, $Z_o(s,t)$ is the cosine of the angle between the surface normal and the Z_o axis (line of sight). Thus, $Z_o(s,t)EGIM(s,t)$ denotes the projected area on the base plane. Since the image plane is parallel to the base plane, the numerator represents the projected area on the image plane. Fig. 4 shows this relationship between the surface normal and the image plane.

The projection area ratio removes only one degree of freedom similar to the reflectance ratio [11]. Fig.5 shows a series of iso-projection-ratio contours of an ellipsoid projected on the stereographic plane [11] for graphical clearness. Each contour line represents possible lines of sight which would give the same projection

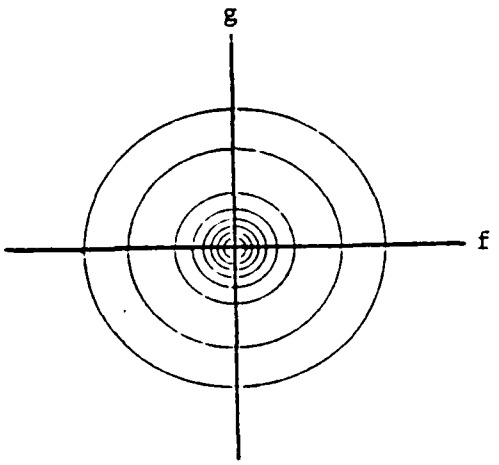


Figure 5. The projection area ratio. Iso-projection-ratio contours of the ellipsoid are drawn on the stereographic plane. Comparing an observed ratio with this diagram prunes the freedom of lines of sight.

---

ratio. The ellipsoid is $X_p^2 + Y_p^2 + (\frac{Z_p}{5})^2 = 1$. The origin of the stereographic plane is aligned with the Z_p axis direction. The stereographic plane axes, f, g agree with the X_p, Y_p axes, respectively. Observing the ellipsoid from the long axis direction gives a smaller projected area than looking at it from the direction perpendicular to the axis. Yet, the surface area is the same. Thus, points near the origin, which correspond to the direction near the long axis, give smaller ratios than far points on the stereographic plane. Therefore, iso-ratio contours are concentric circles around the origin of the stereographic plane.

3.3. EGI inertia direction

The Gaussian hemisphere can also be rotated about the candidate line of sight. This degree of freedom is determined using the 2D EGI inertia axis. This inertia axis is defined on the tangential plane (image plane) to the visible hemisphere at the north pole. Although, it is possible to use the EGI mass center position for this purpose, we prefer the EGI inertia direction for the following reason. The least inertia direction can be determined in any distribution. On the other hand, the EGI mass center position could occur along the line of sight, for example, if the distribution is symmetric with respect to both the X_o, Y_o axes; in this case, the

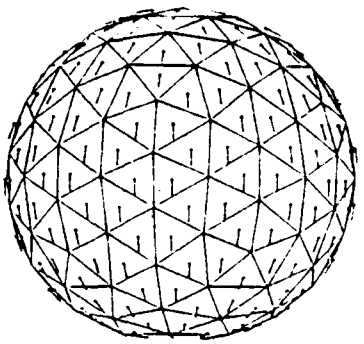


Figure 6. An example of EGI inertia direction. This direction constrains the freedom around the line of sight.

---

alignment cannot be determined. It is true that if an EGI distribution is rotationally symmetric, then we cannot determine the least inertia direction either. However,we need not worry about rotation for rotationally symmetric EGI distributions.

The following calculation gives this axis direction.

$$I_{xx}(v) = \int\int_{V.H.} EGIM(s,t)X_o(s,t)X_o(s,t)\sqrt{EG - F^2}dsdt \tag{4.1}$$

$$I_{xy}(v) = \int\int_{V.H.} EGIM(s,t)X_o(s,t)Y_o(s,t)\sqrt{EG - F^2}dsdt \tag{4.2}$$

$$I_{yy}(v) = \int\int_{V.H.} EGIM(s,t)Y_o(s,t)Y_o(s,t)\sqrt{EG - F^2}dsdt \tag{4.3}$$

$I_{xx}(v), I_{xy}(v), I_{yy}(v)$ gives the principal inertia direction for the line of sight v,

$$\alpha(v) = (\frac{1}{2})\tan^{-1}\frac{2I_{xy}(v)}{I_{xx}(v) - I_{yy}(v)} \tag{5}$$

Thus, $\alpha(v)$ gives the direction of the minimum inertia axis on the image plane.

For example, Fig.6 shows the axis directions of the ellipsoid. A needle depicts the minimum inertia direction for each line of sight. Since this axis direction is unique for each line of sight, this axis constrains the degree of freedom around the direction completely. The resulting direction thus indicates the way a prototype should be aligned with respect to the observed EGI.

4. Implementation of EGI

4.1. Tessellation of the Gaussian sphere

In order to represent and manipulate an EGI on a computer, one must tessellate the Gaussian sphere uniformly. A continuous surface such as an elliptic surface is mapped to a continuous EGI mass distribution. A tessellated sphere is needed to represent this image in a computer. The tessellation method must provide a uniform division of the Gaussian sphere. Since we cannot predict the line of sight, a tessellation method should have the same angular resolution in every direction. Thus, each cell on the tessellated sphere is required to have the same area and the same distance from its adjacent cells.

The projection of a regular polyhedron onto a sphere has this property [31]. A regular polyhedron has faces of equal area, which are evenly distributed in direction with respect to the center of gravity. Thus, projecting edges of a polyhedron onto the circumscribed sphere with respect to the sphere center tessellates the sphere uniformly.

Since the highest order regular polyhedron is the icosahedron, we have to use a geodesic dome [32]. A geodesic dome is obtained by division of each triangle of the tessellated sphere into smaller triangles. We use a geodesic dome from a two frequency dodecahedron, because the geodesic dome has a more uniform facet area distribution than other domes of the same tessellation order [31].

Our computer representation of the tessellated dome has a hierarchical structure. Each cell on one level contains both pointers to sub-cells on the next level and the direction of the center point of the cell. Since we chose a two frequency dodecahedron, the top level cell contains pointers to twelve cells from a dodecahedron. The second level has 60 triangular cells from a one frequency dodecahedron. The lowest level contains 240 triangular cells from a two frequency dodecahedron (See Fig. 7.). The data structure also maintains distance measures between neighboring cells.

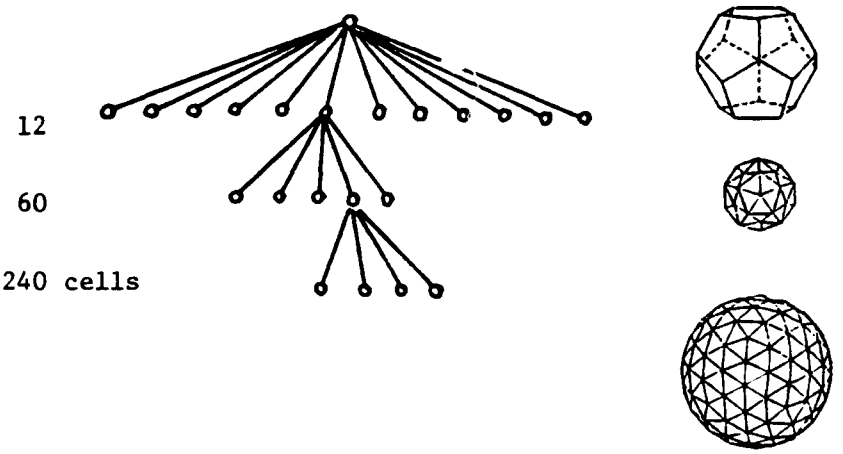


Figure 7. The tessellated dome. The tessellated dome contains 240 triangular cells from a two frequency dodecahedron.

The tessellated dome is used for two purposes. One is to accumulate an EGI image. A particular object surface patch corresponds to a cell with a given surface orientation on the dome. Measured surface area will be added to the corresponding cell. The cumulative image on the dome is the distributed version of the object's EGI. The other purpose is to sample the possible line of sight. Since the cells are distributed uniformly over the dome, the center position of each cell can define the spatial direction. Therefore, the line of sight space is sampled uniformly by this dome.

4.2. Normalized EGI

In the case of a convex object, we only need to store the complete EGI image over the Gaussian sphere in a computer. Since every surface patch whose orientation belongs to the visible hemisphere of v is always observed from the direction v, the EGI obtained from the direction v is exactly the same as the half image of the complete EGI over the hemisphere. Thus, we can derive an EGI over a visible hemisphere by rotating the complete EGI and taking the EGI image over the upper hemisphere. Fig 8 shows some EGIs of convex objects.

In the case of a non-convex object, a surface patch, whose orientation belongs to the visible hemisphere may be is hidden by another part of the object. This

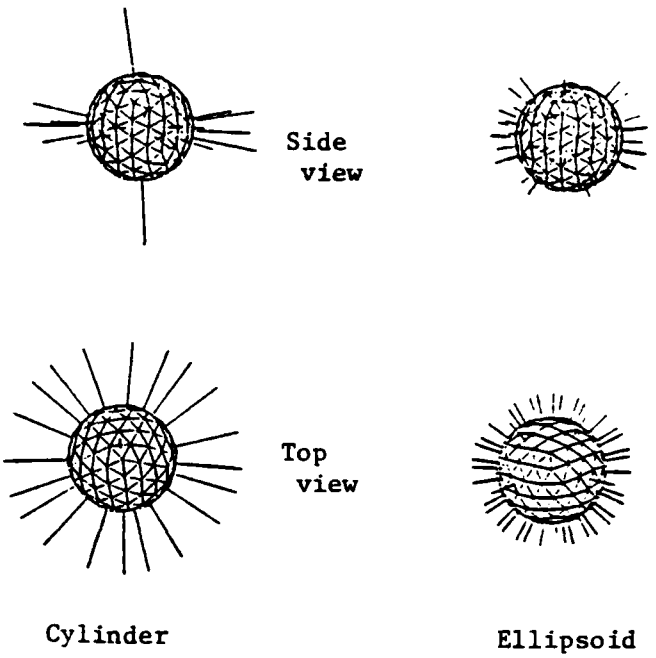


Figure 8. EGIs of convex objects.

occlusion problem requires us to recalculate the EGI for each line of sight. This can be done using either a geometrical modeler [33] or a mathematical expression of the object.

The EGI of a non-convex object can be expressed using four parameters. The line of sight can be expressed using two parameters. A EGI mass distribution at a line of sight is expressed using another two parameters. Namely,

$$EGI == EGIM(s, t, s_l, t_l), \tag{6}$$

where (s_l, t_l) denotes the line of sight expressed on the prototype Gaussian sphere X_p, Y_p, Z_p. (s, t) denotes a point on the visible hemisphere defined by (s_l, t_l). Note that (s_l, t_l) is similar to the light source direction and (s, t) is similar to the surface orientation of the reflectance map [11].

We can store this four dimensional EGI distribution in a two dimensional table. Since tessellation cells on the dome can be ordered along a one dimensional row, an EGI mass distribution for the line of sight v, can be represented as a one dimensional

vector. The possible lines of sight are also ordered in one dimension. Therefore, an EGI can be stored in a two dimensional table, with each row corresponding one possible line of sight. Each element contains an EGI mass (surface area) corresponding to the surface orientation for that line of sight. Since (X_o, Y_o, Z_o) is defined for each line of sight, it is convenient for matching to define the orientation of each column at each row based on each (X_o, Y_o, Z_o).

A discrete version of the projection ratio $A(v)$ for the line of sight v, is obtained by,

$$A(v) = \frac{\sum_{i=1}^{n}[Z_o(i) \cdot EGIM(v,i)]}{\sum_{i=1}^{n} EGIM(v,i)},$$

where n is the total cell number of the v row. $EGIM(v,i)$ is an EGI mass of (v,i) component. $Z_o(i)$ is the Z_o coordinate of i column. Note that (X_o, Y_o, Z_o) is defined at each line of sight.

The following calculation gives this axis direction at the line of sight v.

$$I_{xx}(v) = \sum_{i=1,n} EGIM(v,i)X_o(i)X_o(i) \tag{8.1}$$

$$I_{xy}(v) = \sum_{i=1,n} EGIM(v,i)X_o(i)Y_o(i) \tag{8.2}$$

$$I_{yy}(v) = \sum_{i=1,n} EGIM(v,i)Y_o(i)Y_o(i) \tag{8.3}$$

From these values of $I_{xx}(v)$,$I_{xy}(v)$,$I_{yy}(v)$, the principal inertia direction is obtained in the same manner as for the continuous case.

$$\alpha(v) = (\frac{1}{2}) \tan^{-1} \frac{2I_{xy}(v)}{I_{xx}(v) - I_{yy}(v)} \tag{9}$$

This $\alpha(v)$ gives the direction of the minimum inertia direction.

Storing constraint information adds two additional columns for each row. The first column keeps the projection ratio. The second column stores the original inertia direction relative to (X_o, Y_o). The EGI mass distribution over the remaining elements is rotated so as to agree with the X_o axis. We will refer to this recalculated EGI as the normalized EGI (NEGI). Comparing NEGI from an observed needle

map with NEGI on the table need not to worry about the freedom of rotation around the line of sight.

4.3. Extension for Partial Observation

Real applications often require us to determine surface orientation from a partial region of the hemisphere due to occlusion or some characteristics of the input device. For example, the photometric stereo system can determine surface orientation only where all of the three light sources cast their rays directly. Thus, it is necessary to extend the above mentioned constraints to be applicable to the case of partial observation.

Eg.(3) integrates the EGI mass over the visible hemisphere. Taking integration only over a certain part of the visible hemisphere defines the projection ratio on that partial area. This partial area can also define the least inertia direction and the NEGI there. In other words, all projection ratios, EGI inertia directions, and NEGI's are functions of the area where the EGI mass is integrated.

Let us define a circular region on the Gaussian sphere as an integration area and refer to it as a visible disk. The center of the circle is located at the line of sight. The radius of the visible disk is called the prospect angle ω. The visible hemisphere is a special case of the visible disk which has prospect angle $\omega = \frac{\pi}{2}$. Taking $\omega = \frac{\pi}{2}$ gives the projection ratio of Eg(3). On the other hand, when ω approaches zero, the projection ratio is idential to a delta function which is infinite at the point where surface normal exists, and zero elsewhere. We can define various constraints on various visible disks between these two extreme cases. For example, Fig.9 shows how the iso projection-ratio contours of an ellipsoid $X_p^2 + (\frac{Y_p}{5})^2 + (\frac{Z_p}{10})^2 = 1$ vary with prospect angle. Line of sight directions are expressed on the stereographic plane, where the f,g axis coincides with the x,y axis of the ellipsoid. The r axis denotes the projection ratio at the direction (f, g).

So far we have defined the projection ratio, the least inertia axis, and NEGI for visible disks. We will expand the 2D EGI lookup table into a 3D table. As mentioned above, the first dimension denotes lines of sight. The second dimension corresponds to surface normal directions. And the third dimension corresponds to

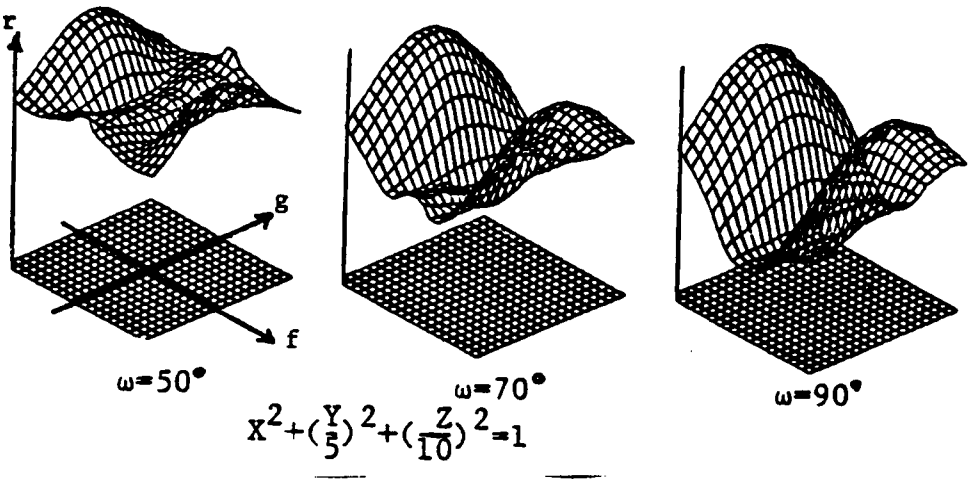


Figure 9. Variation of iso-projection-ratios due to the prospect angle.

prospect angles. This third dimension does not need to have as fine a mesh as the other two dimensions; usually, five or ten tessellating are enough. This 3D table, thus, looks up the projection ratio axis direction, and NEGI, on a certain line of sight over a certain visible disk.

The visible disk can be found from an observed needle map. At first, an observed needle map is converted into a unnormalized EGI on the Gaussian sphere. We can determine an inscribed circle of the obtained EGI distribution. This inscribed circle determines a visible disk which has smaller radius and whose prospect angle is contained in the table. The circle center determines a new pseudo line of sight. The radius determines the observed prospect angle. EGI mass distribution is normalized so as to have unit total mass and the least inertia axis aligned with the X axis over this visible disk. The NEGI at the angle from the table will be compared with the obtained NEGI.

This method works for the following partial observation: (i) An imaging system can determine surface orientation partially. (ii) A curved convex object is occluded by another object. Even in the case that we can observe the whole visible hemisphere, it is better to use NEGI on visible disks of the prospect angle 70 or 80 degrees, for accuracy.

5. Matching Function

A matching function determines a similarity measure between an observed NEGI and the NEGI table. The matching function checks whether each column has a similar amount of EGI mass to the corresponding table column. More precisely, the following operation is done at each column inside of the visible disk. A cumulative sum represents the similarity of row v. Note that this operation is done only when row v's projection area ratio is similar to the observed ratio.

if $EGIM^{observe}(i) = 0.0$

do nothing

if $EGIM^{observe}(i) \neq 0.0$

$\Delta = \frac{|EGIM^{observe}(i) - EGIM^{model}(v, i+e)|}{EGIM^{observe}(i)}$

$d = distance(i, i+e)$

if $d < d^{standard}$, *and* $\Delta < \Delta^{standard}$

add $EGIM^{observe}(i) \cdot (1 - \Delta) \cdot \cos d$ *to total point,* $S(v)$.

$EGIM^{observe}(i)$ is observed NEGI mass at i cell, $EGIM^{model}(v, i+e)$ is EGI mass at $(v, i+e)$ cell of the table, and $distance(i, i+e)$ is the inner product between the cell directions. Δ is the relative error of $EGIM^{model}(v, i+e)$ assumed to correspond to $EGIM^{observe}(i)$. Thus, the first term represents how important the $EGIM^{model}(i)$ is. The second term represents how different the two mass is. The third term represent how far the two mass is. If $EGIM^{model}(v, i)$ has the exactly same mass as $EGIM^{observe}(i)$, $(1 - \Delta) = 1$ and $d = 1$. Then, $EGIM^{observe}(i)$ is added to the total point. If this correspondence is established at each column i, the total point becomes 1, because total EGI mass is 1. This value is the highest score this matching function gives.

$EGIM^{observe}(i)$ are compared not in the column order, i but in nonincreasing order of EGI mass. In order to avoid one prototype cell being put in correspondence with more than one observed cell, a part of the prototype mass at a cell which passes the similarity check is discarded in the amount of the corresponding observed mass.

The direction having the highest score is determined as the observed line of sight. The lookup table also registers how many degrees the prototype is rotated so as to align the least inertia axis with the X axis. Obviously, we know how many degrees the observed image is rotated so as the least inertia axis into coincidence with the X axis. These two angles gives the rotation angle of the observed image around the line of sight.

6. Experiments

6.1. Synthesized image

Consider the simple example of an ellipsoid $X_p^2 + Y_p^2 + (\frac{Z_p}{5})^2 = 1$ observed from the direction inclined from the Z_p axis by an angle of 10 degrees. The surface normal will be derived analytically and used to generate a needle diagram. The algorithm will be applied to the synthesized needle diagram. The result will be used as a basis for judging the performance of the algorithm.

In order to know the performance of the NEGI matching process, we will not use the projection ratio constraint. We will measure the similarity at each 10 degree interval from 0 through 90 degrees. The prospect angle is assumed to be 60 degrees.

A 40 x 40 synthesized needle-diagram was generated. At each image point the outer surface normal of the ellipsoid was calculated. For graphical clearness, each surface normal is depicted as a needle shown in Fig. 10.

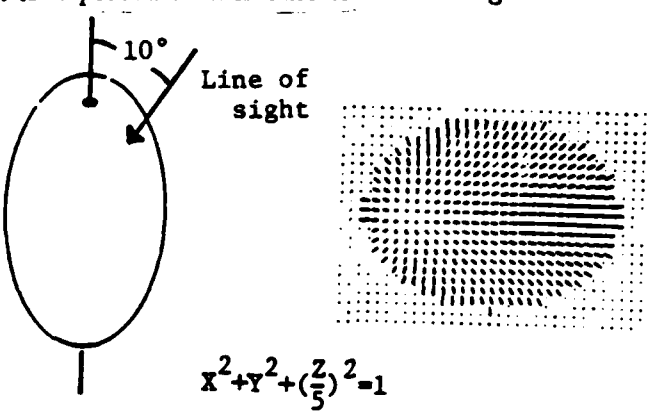


Figure 10. A synthesized needle diagram.

The first task is to obtain an NEGI from this needle diagram. The surface normal at each image point determines the corresponding cell on the Gaussian sphere. Then, EGI mass $\frac{1}{\cos\gamma}$ is added to the cell's EGI mass, where γ is the angle between the surface normal and the line of sight direction. The distribution of EGI mass defines a visible disk. The cumulative mass over the disk is normalized so as to have a unit total mass. Note that both deriving the least inertia axis and normalization is done within the visible disk whose prospect angle is 60 degrees. Fig.11 shows the NEGI obtained from the needle map.

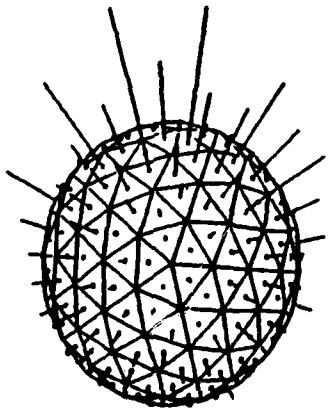


Figure 11. NEGI obtained from the needle diagram.

We will compare the NEGI with prototypical NEGI's registered in the lookup table. Fig. 12 shows some NEGIs in the table. Similarity scores are obtained using the matching function. As indicated in Fig.13, the similarity has its maximum value at 10 degree.

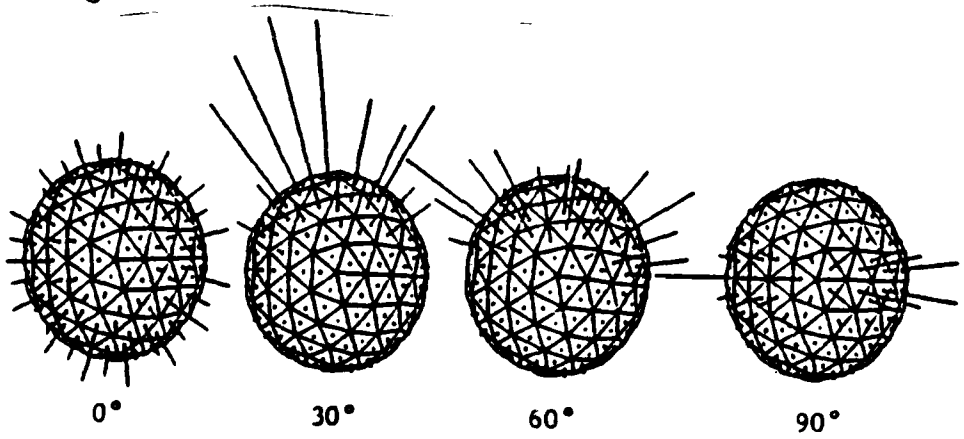


Figure 12. Some NEGIs in the lookup table.

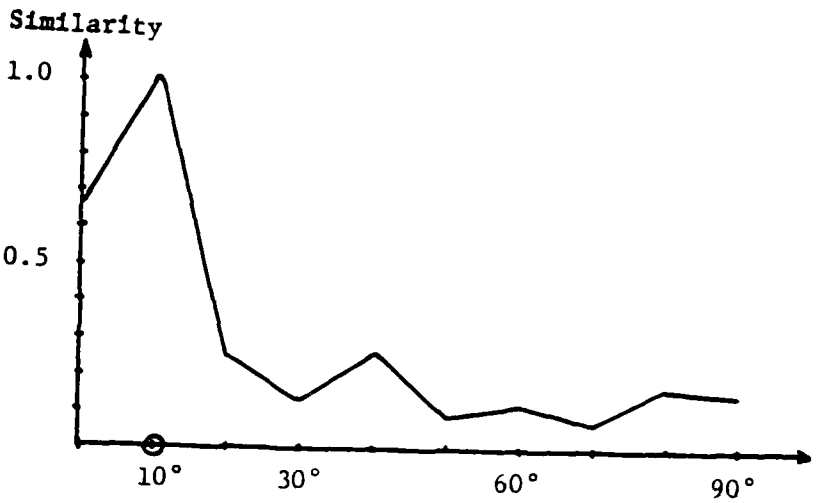


Figure 13. Similarity scores over the candidate directions. The similarity change has its maximum value 1.0 at 10 degree.

6.2. Donuts Experiment

The algorithm is applied to the scene shown in Fig.14 for visual guidance of a manipulator. Note that a simple silhouette matching does not work well in this case.

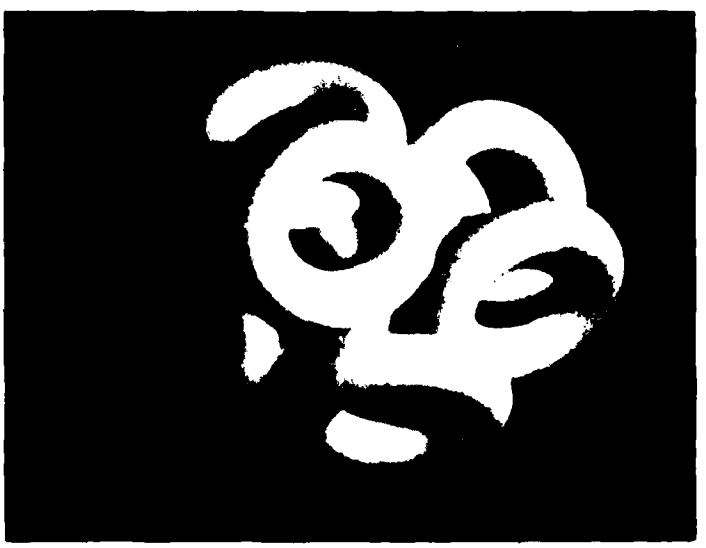


Figure 14. Input scene (a pile of objects).

The donuts needle diagram is obtained using the photometric stereo system [14-16]. Fig.15 shows the needle diagram obtained. The photometric stereo system can determine surface orientations for areas whose prospect angle is no more than 50 degrees.

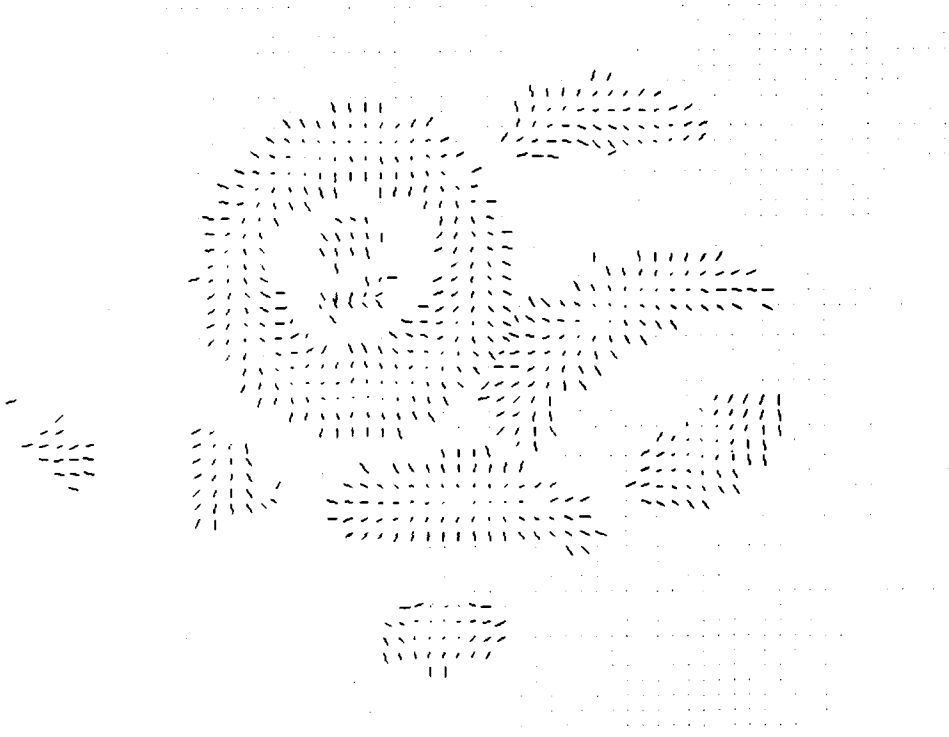


Figure 15. The needle diagram obtained by the photometric stereo system.

A segmentation program based on surface smoothness is applied and a target area is selected [34]. The region is bounded not with occluding boundaries but with internal curves. Fig.16 shows the segmented regions. Fig. 17 shows the target region selected by a decision process [34].

The projection ratio, the least inertia axis, and the NEGI are calculated from surface normals at the target region. Fig.18. shows the obtained NEGI.

Since a donut has a rotational symmetry axis, the necessary sampling directions are points along a 90 degrees section of a great circle on the Gaussian sphere

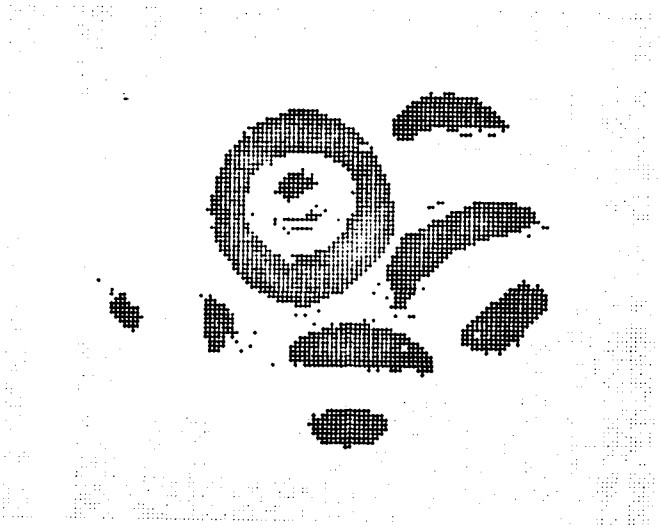


Figure 16. The segmented regions.

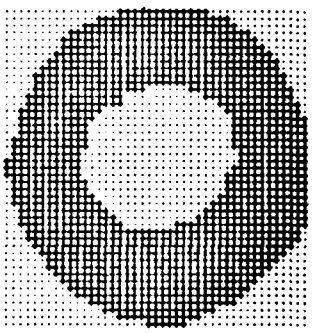


Figure 17. The target region selected.

containing the symmetry axis. At 10 degree increments along the great circle, the NEGI, projection ratio, and the least inertia axis direction are calculated using a mathematical model of the donut shape. Since we know that the prospect angle is 50 degrees, only one layer of the 3D lookup table is calculated. Since directions near the axis may have relatively large error in the least inertia axis direction, NEGI's rotated by a small amount around the exact alignment are also registered in the lookup table. Fig.19 shows the output of the algorithm which indicates the approach direction of the end effector of the arm.

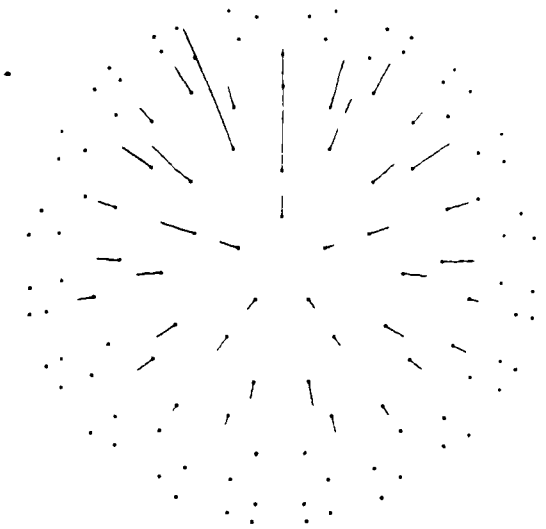


Figure 18. NEGI obtained from the target region.

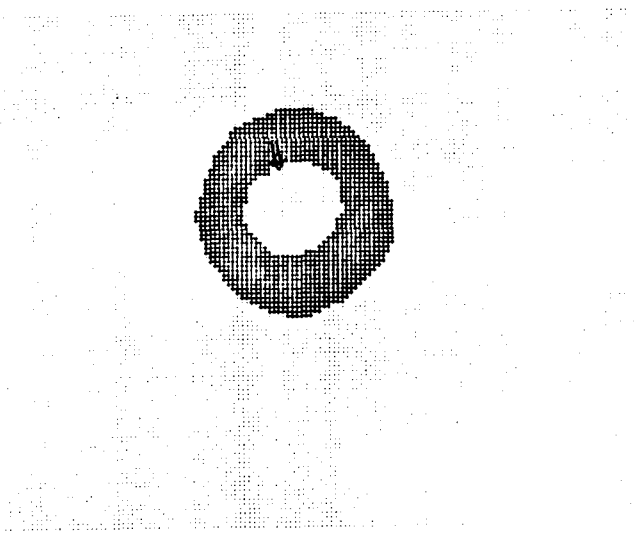


Figure 19. An approach direction of an end effector.

7. Concluding Remarks

This paper determines the attitude of an object using an extended Gaussian image. The attitude of an object has three degrees of freedom. The freedom is greatly reduced if we apply constraints derived from a global distribution of EGI

mass. One constraint comes from EGI mass center position. The other constraint is based on the least EGI mass inertia direction. After reducing the attitude possibilities with these constraints, a final decision is made comparing it with the model's EGIs in certified attitude. The best fitting attitude is selected as the observed attitude of the object.

In case that there exist a various kinds of object in a scene, we can make an extended NEGI table joining each NEGI table of each object. The matching process may follow the same procedure described above. Usually the EGI matching can determine the most likely attitude of the most likely object. In the worst case, however, since EGI matching is a necessary condition for congruence of general objects, we have to examine position information to identify the object. In this case a decision process can make a final decision comparing the observed position information with the information produced from a geometical modeler [33] at the determined attitude [1]. Since the attitude is already determined, the calculation cost of the geometrical modeler is cheap.

Acknowledgements

The author would like to extend his sincere appreciation to Yoshiaki Shirai and Berthold Horn for their helpful discussions. Kokichi Sugihara, Berthold Horn, Mike Brady, Eric Grimson, Keith Nishihara, and Brian Schunck proofread drafts of this article. Their comments are appreciated. John Canny provids us array-operation libraries microcoded on Lispmachines. Thanks go to Ikko for preparing this manuscript and drawings. This research is done in part at Electro-Technical Laboratory, MITI, and in part Artificial Intelligence Laboratory, MIT.

References

[1] Ikeuchi, K. and Shirai, Y., "A Model Based Vision System for Recognition of Machine Parts," *Proc AAAI-82*, Pittsburg, 18-21.

[2] Price, K. and Reddy, R., "Matching Segments of Image," *IEEE Trans. on PAMI* Vol. PAMI-1, No. 1, Jan. 1979, 110-116.

[3] Davis, L.S., "Shape Matching Using Relaxation Techniques," *IEEE Trans. on PAMI*, Vol. PAMI-1, No. 1, Jan. 1979, 60-72.

[4] Sakane, S., "An Algorithm for Scale Invariant Segment-Matching," *Proc. PRIP-81*, Dallas, 565-571.

[5] Brooks, R., "Symbolic Reasoning among 3-D Models and 2-D Images," *Artificial Intellignece*, Vol 17, August, 1981, 285-349.

[6] Marr, D., *Vision* W.H. Freeman, San Francisco, 1982.

[7] Brady, M., "Computaional Approaches to Image Understanding," *ACM Computing Surveys*, Vol 14, 1982, 3-71.

[8] Ballard, D.H. and Brown, C.M., *Computer Vision*, Prentice-Hall, New Jersey, 1982.

[9] Grimson, W.E.L., *From Images to Surface: A Computational Study of the Human Early Visual System*, MIT Press, Cambridge, 1981.

[10] Horn, B.K.P., "Obtaining Shape from Shading Information," *The Psychology of Computer Vision*, P.H. Winston, ed., 115-155, McGraw-Hill, New York, 1975.

[11] Ikeuchi, K and Horn, B.K.P., "Numerical Shape from Shading and Occluding Boundaries." *Artificial Intelligence* , Vol. 17 1981.

[12] Brooks, M.J., "Shape from Shading Discretely," *Ph.D. Thesis*, Essex University, Sept 1982.

[13] Terzopoulos, D., "Multi-level Reconstruction of Visual Surfaces" *AI Memo. 671*, AI Lab., MIT, 1982.

[14] Woodham, R.J., "Reflectance Map Techniques for Analyzing Surface Defects in Metal Castings," *AI-TR-457*, AI Lab., MIT, June 1978.

[15] Ikeuchi, K., "Determining Surface Orientations of Secular Surfaces by Using the Photometric Stereo Method," *IEEE Trans. on PAMI* , Vol. PAMI-2, No. 6, Nov. 1981, 661-669.

[16] Coleman, E.N. and Jain, R.,"Shape from Shading for Surfaces with Texture and Specularity," *Proc. IJCAI-81*, Vancouver, August, 1981, 652-657.

[17] Kender, J.R., "Shape from Texture," *Ph.D. Thesis* , Computer Science Dept., CMU, 1980.

[18] Kender, J.R., and Kanade, T., "Mapping Image Properties into Shape Constraints," *Proc. AAAI-80,* Stanford, August, 1980, 4-6.

[19] Ikeuchi, K., "Shape from Regular-patterns An Example of Constraint Propagation in Vision." in *Proc. IJCPR-80* , Miami Beach, Dec., 1980, 1032-1039.

[20] Ohta, Y., Maenobu, K., and Sakai, T., "Obtaining Surface Orientation from Texels under Perspective Projection," *Proc IJCAI-81*, Vancouver, August, 1981, 746-751.

[21] Witkin, A.P. "Recovering Surface Shape and Orientation from texture," *Artificial Intellignec*, Vol. 17, 1982, 17-47.

[22] Horn, B.K.P., "SEQUINS and QUILLS – representation for surface tomography." *AI Memo No. 536*, AI Lab., MIT, 1979.

[23] Barrow, H.G. and Tenenbaum, J.M., "Recovering Intrinsic Scene Characteristics from Images." Hanson, A.R. —& Riseman, E.M., (Ed.) *Computer Vision Systems* , Academic Press, New York, 1978.

[24] Marr, D. and Nishihara, K.H., "Representation and Recognition of the Spatial Organization of Three-dimensional Shape." *Proc. R. Soc. Lond.* , Vol. B:200, 1978.

[25] Binford, T.O.,"Visual Perception by Computer." *IEEE Conf. on Systems and Control* , Miami, Dec. 1971.

[26] Smith, D.A.,"Using Enhanced Spherical Images for Object Representation," *AI Memo No. 530*, AI Lab., MIT, 1979.

[27] Bajcsy, R.,"Three-dimensional Scene Analysis," in *Proc. ICPR-80* , Miami Beach, Dec. 1980, 1064-1074.

[28] Ballard, D.H. and Sabbah, D., "On Shapes," *Proc. IJCAI-81*, Vancouver, August, 1981, 607-612.

[29] Brou, P., "Finding the Orientation of Objects in Vectors Maps," *Ph.D. Thesis*, EECS, MIT, (in preparation).

[30] Horn, B.K.P., A Personal Communication, Dec. 1980.

[31] Wenninger, M.J., *Spherical Models* , Cambridge University Press, Cambridge, 1979.

[32] Hosaka,M., Kimura, F., and Kakishita, N., "A Unified Method for Processing Polyhedra," *Proc. IFIP-74*, 768-772.

[33] Ikeuchi, K., Horn, B.K.P., and Nagata, S. "Picking up an Object from a Pile of Object," *Proc. Int. Sym. Robotics Research,* Cambridge, September, 1983, (in preparation).